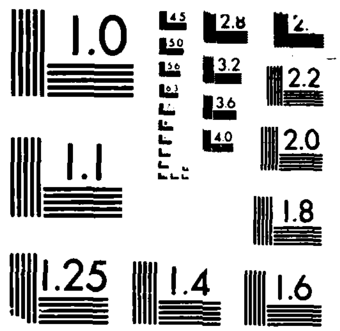
AD-A193 692 HAPTIC EXPLORATION IN HUMANS AND MACHINES: ATTRIBUTE INTEGRATION AND MACH..(U) CALIFORNIA UNIV SANTA BARBARA COMMUNITY AND ORGANIZATION RESE.. S J LEDERMAN ET AL. 30 APR 88 TR-88-01 N00014-86-K-0232 1/1

UNCLASSIFIED F/G 23/2 NL



MICROCOPY RESOLUTION TEST CHART

IREAU ... STANDARDS 1963 A

AD-A193 692

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: 88-01	2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER	DTIC FILE COPY
4. TITLE (and Subtitle): HAPTIC EXPLORATION IN HUMANS AND MACHINES: ATTRIBUTE INTEGRATION AND MACHINE RECOGNITION/IMPLEMENTATION	5. TYPE OF REPORT & PERIOD COVERED: Technical Report May 1, 1987 - April 30, 1988
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): R.L. Klatzky, S.J. Lederman, R. Bajcsy, S. Stansfield, N. Ulrich, & R.P.Paul	8. CONTRACT OR GRANT NUMBER(s): N00014-86-K-0232
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Community and Organization Research Institute, University of California, Santa Barbara, CA 93106	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: R&T4424200
11. CONTROLLING OFFICE NAME AND ADDRESS: Engineering Psychology, Office of Naval Research, Arlington, Virginia 22217	12. REPORT DATE: April 1, 1988
	13. NUMBER OF PAGES: 54
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS. (of this report): Unclassified
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

DTIC SELECTE D APR 25 1988 S E

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

haptic, hand, touch, vision, robot, object recognition, categorization

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This technical report comprises three documents. The first describes the integration of haptic attributes during object categorization. The second describes a machine object-recognition system with haptic as well as visual sensors. The third describes the development of a novel end effector of medium complexity.

DD FORM 1473, 1 JAN 73 EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-LF-014-6601

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

HAPTIC EXPLORATION IN HUMANS AND MACHINES:

ATTRIBUTE INTEGRATION AND MACHINE RECOGNITION/IMPLEMENTATION

This technical report comprises three documents. The first describes the integration of haptic attributes during object categorization. The second describes a machine object-recognition system with haptic as well as visual sensors. The third describes the development of a novel end effector of medium complexity.

The first document was presented as a report to the annual meeting of the Psychonomic Society, 1987. Documents 2 and 3 were issued as technical reports from the University of Pennsylvania Dept. of Computer and Information Science (#s MS-CIS-87-61 and MS-CIS-87-82).

Accession For	
NTIS GRA&I	X
DTIC TAB	☐
Unannounced	☐
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	

88 4 25 151

Haptic Categorization of Objects by Multiple Dimensions

Roberta L. Klatzky — Univ. of Calif. Santa Barbara
Susan Lederman — Queen's University Kingston, Ont.
Catherine Reed — Univ. of Calif. Santa Barbara

(presented at the 1987 meeting of the Psychonomic Society)

Our previous work, much of which has been reported at past Psychomic Society meetings, has established that the haptic system has remarkable capabilities for object recognition. We define haptics as purposive touch. The basic tactual system incorporates information from cutaneous sensors in the skin and kinesthetic sensors in muscles, tendons, and joints. Its sensory primitives therefore include pressure, vibration, position, and thermal properties. We have argued, however, that the functional sensitivities of haptics are considerably enhanced by the execution of stereotyped motor patterns, which we call "exploratory procedures" (Klatzky & Lederman, 1987; Lederman & Klatzky, 1987). An exploratory procedure is a motor activity that is typically used for extracting a particular object property. In previous work , we have described the links between desired knowledge about object properties and the nature of exploratory procedures. We have also shown that the procedure that is typically performed to extract a property is generally the optimal one, in terms of accuracy and/or speed.

The procedures we have studied are shown on the first slide. They are lateral motion (a rubbing like action) for encoding texture; pressure for encoding hardness; static contact for thermal sensing; unsupported holding for weight; enclosing for volume and gross contour information; and contour following, which is used to extract precise contour information as well as global shape. We have also considered procedures for encoding higher-level object properties, such as functional uses based on structure, and the nature of part motion.

SLIDE 1 HERE

Although we can distinguish among haptically encoded object dimensions and can couple each dimension with particular exploratory motor movements, this does not mean that the haptic system extracts and processes each dimension independently. In the present work, we addressed the issue of how dimensions are processed together. Specifically, we asked whether information about multiple object dimensions is integrated in haptic processing.

Our approach to this issue is most directly related to Garner's (1974) research on the integrality and separability of stimulus dimensions. This work has made extensive use of classification tasks, in which stimuli are to be assigned to distinct categories on the basis of some dimensional value. For example, large stimuli may be in class A and small in class B.

If a second, redundant dimension is added -- for example, all large stimuli are red and all small stimuli are green -- then either dimension -- color or size -- could be the basis for classification. If classification time is reduced under these circumstances, there is said to be a "redundancy gain." On the other hand, there may be an irrelevant dimension that varies orthogonally to the decision -- for example, half of the large stimuli are circles and half squares, and the same distribution holds for the small stimuli. If classification time increases under these circumstances, there is an "orthogonality loss." In general, redundancy gain and orthogonality loss indicate that information from the two manipulated dimensions has been integrated, so that they jointly contribute to classification.

Note that this pattern does not necessarily justify a stronger claim, that the dimensions are "integral." (See Garner, 1974, p. 152, for the distinction between information integration and dimensional integrality.) To be integral, dimensions must be functionally fused in processing, without volitional control.

Our initial hypothesis was that the haptic system would integrate information about two substance dimensions, texture and hardness, more than the combination of either one with a structural dimension, shape or size. There are several reasons for this prediction. First, texture and hardness are both typically extracted by local exploration of a homogeneous object surface. In contrast, shape and size information are extracted through exploration of the outer object envelope, through contour following or enclosure. Although it would be possible to determine texture and hardness information while exploring along a contour, the preference for extracting these dimensions from different parts of objects may mean that haptics does not naturally process structure and substance dimensions together. Moreover, our previous work (Klatzky, Lederman, & Reed, in press) had demonstrated that texture and hardness information are both highly salient to haptic explorers who are learning about an object's properties. Shape was less so, and size was particularly low in salience, although this may reflect the hand-size range of our particular stimuli. The salience effects suggest that the shape and substance dimensions are differentially weighted, if not actually segregated, in object processing. Finally, we have recently gathered ratings of the importance of dimensions for categorizing common objects by touch. Texture and hardness ratings strongly co-vary, which is consistent with the idea that they are integrated in haptic exploration.

In our first experiment, we asked subjects to sort a set of multidimensional stimuli that potentially varied on 4 dimensions --hardness, size, roughness, and shape (as shown on slide 2). There were factorial combinations of 3 values on each dimension. The objects had been constructed so that the single dimensions were all about equally well discriminated. Tests of sorting time along each dimension validated this goal, except for size, which was somewhat less discriminable. Thus we focussed on the

remaining three dimensions -- shape, texture, and hardness -- in the classification task.

SLIDE 2 HERE

Subjects were assigned to 7 groups, according to the following slide. In each of three one-dimensional groups, the classification decision was made on the basis of only one dimension. Each level of this dimension defined a different class. For example, all round objects might be A, all hourglass shapes B, and all clover shapes C. In each of three two-dimensional groups, either of two redundant dimensions was sufficient for classification. And in a three-dimensional group, the three dimensions were redundant indicators of the stimulus class. Note that we covaried redundancy and orthogonality here, to maximize the potential for observing group differences. If a dimension was not redundant, it varied orthogonally to the response decision. (Size varied orthogonally in all conditions, for reasons described above.)

SLIDE 3 HERE

Each blindfolded subject repeatedly classified 9 objects. Subjects were not told what dimension or dimensions was relevant to their partitioning of the stimuli, but they were allowed to explore the stimuli at the beginning of the task, and they were required to correctly classify each one before beginning speeded trials. On each trial, the stimulus was placed on a force-sensitive board with a piezoelectric sensor. The experimenter then readied the computer, which emitted a beep to signal to the subject that the object was in position. Upon first contact with the object, a signal from the board started a clock, which terminated when the subject vocalized the stimulus class. Thus response times were recorded. In addition, we videotaped subjects performing the task and analyzed their hand movements.

The next slide shows the classification time for each group, over a sequence of 144 trials, in 3 blocks. There is an overall practice effect, but more important, there are differences among the groups. The groups with one relevant dimension did not significantly differ, as we expected given our construction of the dimensions to be about equally discriminable. One-dimension classification was slower than two, but three dimensions did not produce a gain over two. Among the two-dimension groups, there was a tendency for texture + hardness to be fastest. (This did not reach significance in these data, but did in the next experiment to be described.)

SLIDE 4 HERE

Why should there be integration of two dimensions, but not three? In answer, we turn to the data on the hand movements of subjects in the various groups. These data consist of the percentage of trials, out of a sample from each period, that demonstrated 4 targeted exploratory procedures: lateral motion for texture, pressure for hardness, and enclosure and contour

following for shape. Considering the two-dimension groups, there was a general tendency for relevant exploratory procedures to emerge at least by the last block of trials. Particularly striking was the pattern for the texture/hardness group, which concentrated exclusively on relevant exploratory procedures from the very beginning. In fact, frequently both of these procedures were used on the same trial, often in the form of a hybrid "smear" that moved across the surface of the object with noticeable normal force. The three-dimension group showed a pattern highly similar to the texture/hardness group; in fact, their percentages of procedure use correlated .90.

SLIDE 5 HERE

These results are generally consistent with our hypothesis that substance-related dimensions would be natural candidates for information integration in haptics. The data suggest that given all three redundant dimensions, exploration for shape is virtually dispensed with, and exploratory procedures for texture and hardness are executed. Accordingly, the redundant shape information adds little; response times show no reduction relative to a condition in which only texture and hardness are relevant to classification. Note that the the two-dimensional conditions combining shape with texture or shape with hardness do show some advantage over one dimension, and exploration for both dimensions does occur.

Essentially, the limitation on information integration here appears to reflect a limitation on the diversity of haptic exploration. Subjects executed two exploratory procedures, when relevant, but not three. The source of this limitation is yet somewhat ambiguous. For one possibility, subjects could elect to execute redundant procedures because they are motorically compatible. For example, texture and hardness are very compatible, being capable of execution in tandem through a pressurized smear. But pressure and contour following are far less so, because pressure may deform an object's contour or may prevent the hand from moving smoothly along the edge. On the other hand, the limitation on exploration may be secondary to cognitive preferences for combining information about object dimensions. If information from two sources is not integrated, there is no reason to explore for both.

Our next experiment used a converging operation to identify dimensions on which information is integrated. We asked whether the withdrawal of a redundant dimension would impair classification performance. Subjects were trained on the classification task with two redundant dimensions. After 108 trials, they were introduced to a new set of 9 stimuli, which were partitioned into classes defined by only one of the previously relevant dimensions. The other dimension was now withdrawn; it was held constant at an arbitrary value. If information from the withdrawn dimension had previously been used to determine classification, we would expect to see an increase in response time. We call this increase the "dimension

withdrawal effect." (We wanted such an increase to be attributable to adjustment of the classification rule. To avoid a spurious increase from motor practice, the first few trials after the shift were discarded.)

If one of the two redundant dimensions dominates classification initially, we should see an asymmetric withdrawal effect: Subjects from whom the dominant dimension is withdrawn should be impaired, but those for whom the dominant dimension remains informative should not be. In contrast, if both dimensions contribute to classification, withdrawal of either should impair performance.

The results are shown in the next slide. The asymmetric pattern that shows dominance by one dimension is shown for the texture/shape and hardness/shape groups. In this case, shape appears to be given higher weight in classification, because its withdrawal produces an increase in response time. In contrast, the texture/hardness groups show the symmetric pattern of impairment that indicates both dimensions contributed to the decision. Withdrawal of either texture or hardness produced a response-time increment. Thus we find additional evidence for integration of information about substance dimensions.

SLIDE 6 HERE

An analysis of hand movements indicated that prior to the shift, subjects were generally using exploratory procedures relevant to both dimensions, in some mixture. When one dimension was withdrawn, however, they promptly shifted away from the corresponding exploratory procedure, concentrating on the relevant one. This suggests that the dimension-withdrawal effect was not due to perseveration on inappropriate motor activity, but rather reflects the need to adjust dimensional processing.

In a third experiment, we asked whether classifiers who were told that one particular dimension was relevant would still gain from having a second redundant dimension. This addresses the issue of whether integration occurs without explicit instruction. We again used the withdrawal paradigm. Subjects were given a series of classification trials with stimuli that could be classified by either of two redundant dimensions. However, they were told in advance to use one particular dimension for the classification decision. After more than 100 trials, the second dimension, about which subjects had not been informed, was switched from redundant variation to no variation -- that is, its value now was held constant. The next slide shows the effects of this manipulation.

SLIDE 7 HERE

There was a very substantial increase in response time immediately after withdrawal of the redundant dimension, for the conditions in which texture covaried with hardness. Whether subjects were initially told to focus on texture or hardness did not significantly alter the shift. The groups for which the

dimension of shape was redundant with a substance dimension, texture or hardness, showed much less effect, which in most cases was not significant. Thus it appears that texture and hardness were integrated even when instructions biased against doing so, whereas there was little integration of shape and substance.

To summarize, we now have multiple lines of evidence for the integration of texture and hardness in haptic classification. In contrast, the integration of shape information with either of these substance dimensions is more limited. When shape is redundant with texture and hardness, the latter two are the preferred sources of information. The combination of texture and hardness leads to fastest classification, and withdrawal of either dimension impairs performance, whether or not subjects are told about the redundancy. Execution of exploratory procedures generally parallels the observed patterns of dimensional integration. In haptics, we might say, "how you touch is what you get."

References

Garner, W. (1974). The processing of information and structure. Hillsdale, N.J.: Erlbaum.

Klatzky, R.L. & Lederman, S.J. (1987). The intelligent hand. In G. Bower (Ed.), The Psychology of Learning and Motivation, vol. 21, pp.121-151.

Klatzky, R.L., Lederman, S., & Reed, C. (1987). There's more to touch than meets the eye. Journal of Experimental Psychology: General, in press.

Lederman, S., & Klatzky, R. (1987). Hand movements: A window into haptic object recognition. Cognitive Psychology, 19, 342-368.

Slides:

Note on Abbreviations
Groups: T = texture, H = hardness, F = form
Exploratory Procedures: LM = lateral motion, EN = enclosure, CF = contour following, PR = pressure).

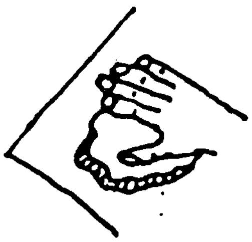
1) Pairing of objects and hand movements.
2) Objects used in study.
3) Nature of Groups, Experiment 1.
4) Response times, Experiment 1, by block and group.
5) Exploratory procedures, Experiment 1, by block and group (2 and 3 dimensions only).
6) Response times, Experiment 2, by period (a.b indicates part a, period b, with 2.1 the point of shift) and group (arrow indicates initially relevant dimensions on left; ultimately relevant dimension on right).
7) Response times, Experiment 3, by period and group (legend as in slide 6).

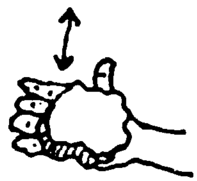
EXPLORATORY PROCEDURE / KNOWLEDGE ABOUT OBJECT

LATERAL MOTION / TEXTURE

PRESSURE / HARDNESS

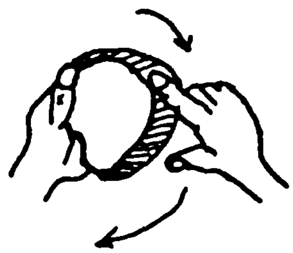
STATIC CONTACT / TEMPERATURE





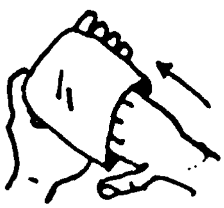
UNSUPPORTED HOLDING / WEIGHT

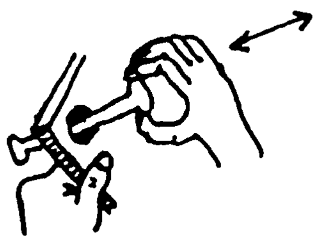
ENCLOSURE / GLOBAL SHAPE, VOLUME



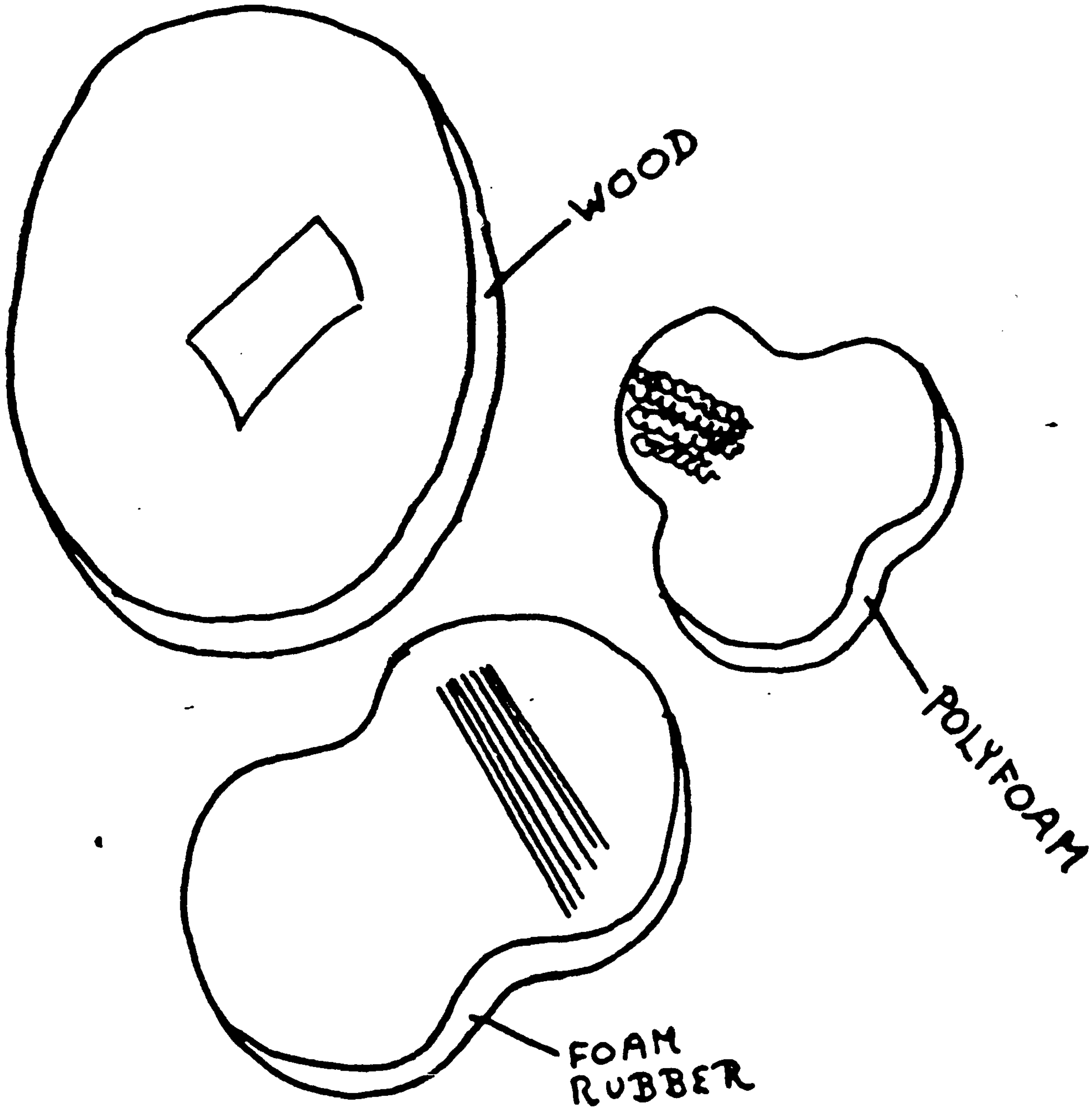
CONTOUR FOLLOWING / GLOBAL SHAPE, EXACT SHAPE

FUNCTION TEST / SPECIFIC FUNCTION





PART MOTION TEST / PART MOTION


WOOD
POLYFOAM
FOAM
RUBBER

GROUPS IN CLASSIFICATION EXPERIMENT

ALL GROUPS: CLASSIFY 9 OBJECTS INTO 3 CATEGORIES (A,B,C)

1. CLASSIFICATION BY HARDNESS ONLY

   Example: A = hard, B = soft, C = medium-hard
   Each class represents all 3 shapes, textures, sizes.

2. CLASSIFICATION BY SHAPE ONLY

3. CLASSIFICATION BY TEXTURE ONLY

4. CLASSIFICATION BY HARDNESS AND SHAPE

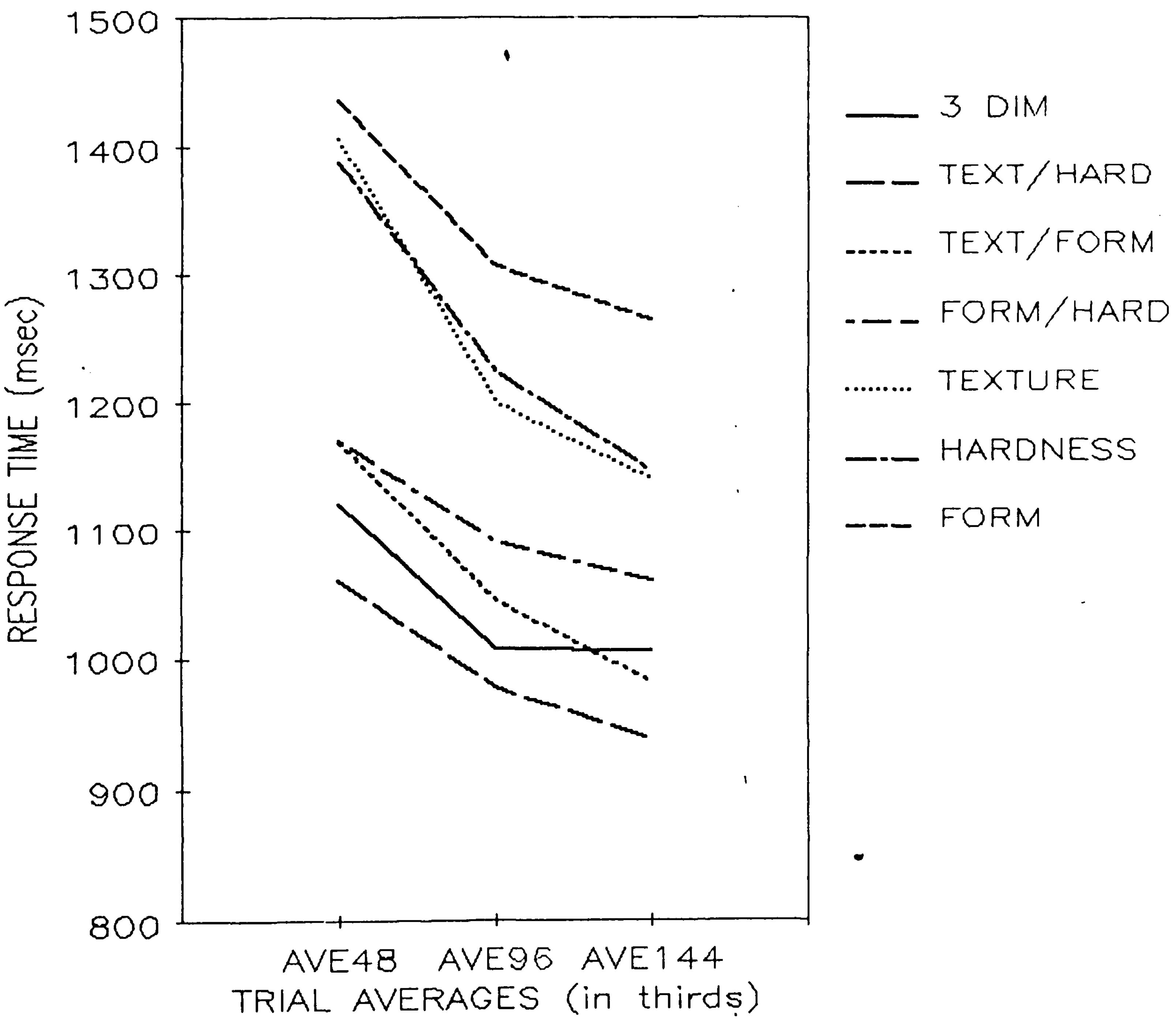
   Example: A = soft oval
   B = medium-hard hourglass
   C = hard clover-shape
   Each class represents all 3 textures, sizes.

5. CLASSIFICATION BY TEXTURE AND SHAPE

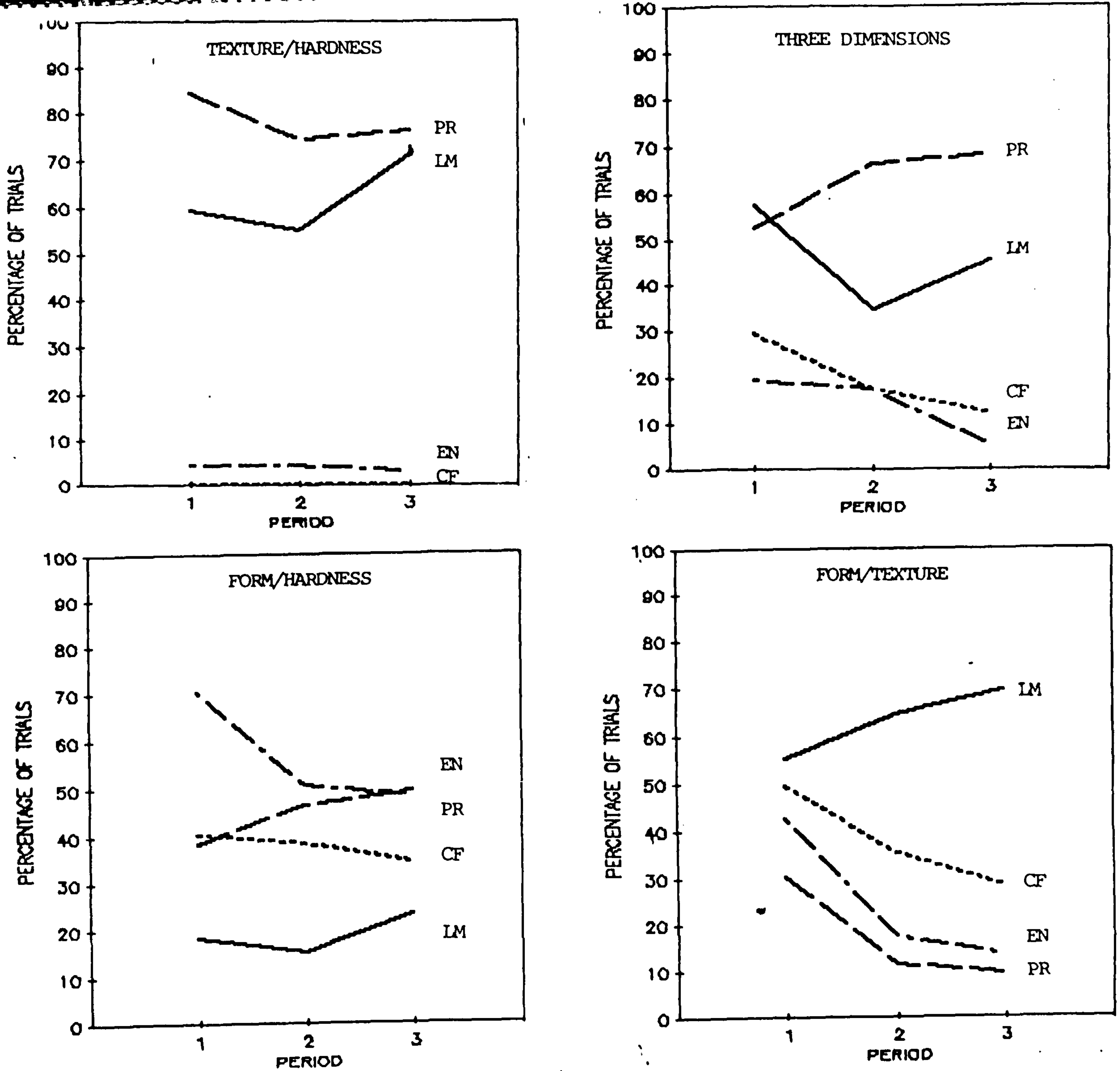
6. CLASSIFICATION BY HARDNESS AND TEXTURE

7. CLASSIFICATION BY HARDNESS, SHAPE, TEXTURE

   Example: A = medium-hard, rough, clover-shape
   B = hard, smooth, hourglass
   C = soft, medium-rough, oval
   Each class represents all 3 sizes.

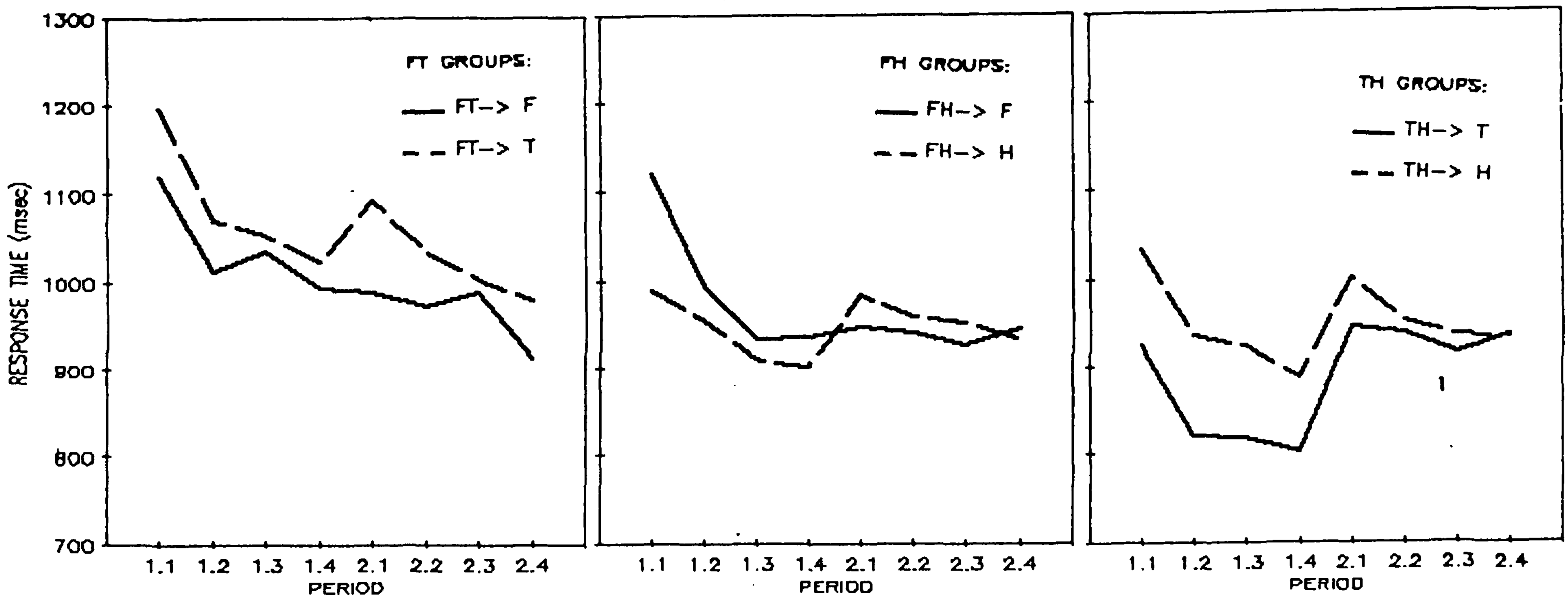


RESPONSE TIME (msec)
1500
1400
1300
1200
1100
1000
900
800
AVE48
AVE96
AVE144
TRIAL AVERAGES (in thirds)
3 DIM
TEXT/HARD
TEXT/FORM
FORM/HARD
TEXTURE
HARDNESS
FORM

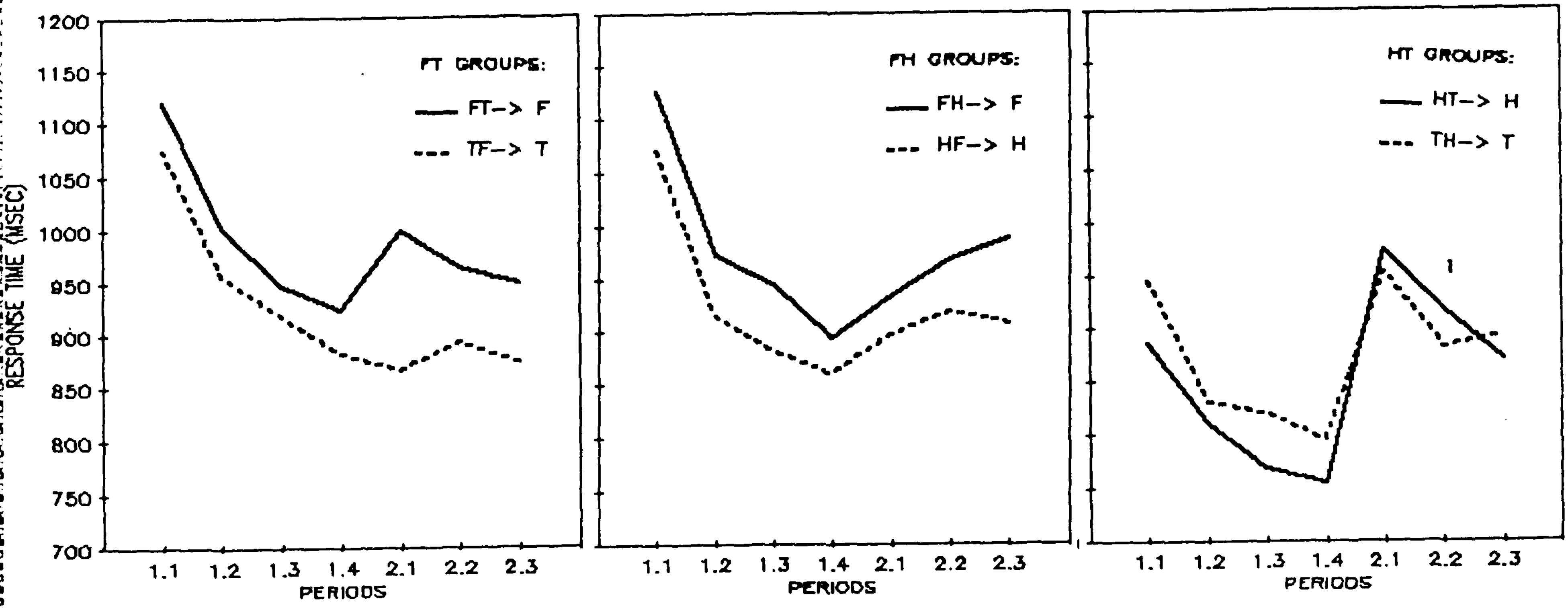


TEXTURE/HARDNESS
PERCENTAGE OF TRIALS
PERIOD
PR
LM
EN
CF
THREE DIMENSIONS
PERCENTAGE OF TRIALS
PERIOD
PR
LM
CF
EN
FORM/HARDNESS
PERCENTAGE OF TRIALS
PERIOD
EN
PR
CF
LM
FORM/TEXTURE
PERCENTAGE OF TRIALS
PERIOD
LM
CF
EN
PR

1



RESPONSE TIME (msec)
1300
1200
1100
1000
900
800
700
FT GROUPS:
FT-> F
FT-> T
FH GROUPS:
FH-> F
FH-> H
TH GROUPS:
TH-> T
TH-> H
1.1 1.2 1.3 1.4 2.1 2.2 2.3 2.4
PERIOD


RESPONSE TIME (MSEC)
1200
1150
1100
1050
1000
950
900
850
800
750
700
FT GROUPS:
FT-> F
TF-> T
FH GROUPS:
FH-> F
HF-> H
HT GROUPS:
HT-> H
TH-> T
1.1 1.2 1.3 1.4 2.1 2.2 2.3
PERIODS

REPRESENTING GENERIC OBJECTS FOR EXPLORATION AND RECOGNITION

Sharon A. Stansfield
MS-CIS-87-61
GRASP LAB 112

Department of Computer and Information Science
School of Engineering and Applied Science
University of Pennsylvania
Philadelphia, PA 19104-6389

June 1987

Acknowledgements: This research was supported in part by DARPA grant DARPA/ONR N0014-85-K-0807, NSF-CER grants MCS-8219196, AVRO DAAB07-84-K-FO77, NIH 1-RO1-HL-29985-01, MCS 82-07294 and U.S. Army grants DAA29-84-K-0061, DAA29-84-9-0027, AfOSR 82-NM-299.

REPRESENTING GENERIC OBJECTS

FOR EXPLORATION AND RECOGNITION

S. A. Stansfield

GRASP Lab, University of Pennsylvania[1]

Philadelphia, Pa. 19104

Abstract

Generic objects are familiar to all of us -- as a matter of fact, we spend our lives surrounded by them. We speak, for instance, of cups and shirts and hammers, usually reverting to more specific descriptions (such as the blue porcelain teacup with the fluted rim) only when it is necessary to distinguish between two objects within the same basic category. It would seem reasonable, then, to give robots this same capability of reasoning in terms of classes of objects. In this paper we present a knowledge representation mechanism for reasoning about generic objects. The task is active tactile exploration for object identification. Objects are first imaged visually and are then explored haptically. Our object representation is feature-based, with geometric/spatial information coming from a model which we call the *spatial polyhedron*. If there is only one hypothesis about the identity of the object, the system generates verification strategies. If there is more than one hypothesis, then the system uses feature-based reasoning to generate strategies for distinguishing among the various possibilities.

1. Introduction

When people speak of cups or screwdrivers, they may or may not have a specific object in mind. If you were asked to take the cup from the baby, you would have no trouble identifying the object in the baby's hands as the desired object (providing the baby was holding only one cup.) Likewise, if someone were to ask you to draw "a cup", you could probably do so without asking which cup they had in mind. Thus people tend to speak, reason, and perceive in terms of generic, rather than specific, objects, reverting to more specific descriptions only when it is necessary to distinguish between two objects within the same class. (In our example, if the baby were holding both a blue, clay mug and a pink, plastic teacup, you might have to ask

[1]The work described herein was partially supported through the following grants: ARO DAA6-29-84-k-0061; AfOSR 82-NM-299; NSF MCS-8219196-CER; NSF MCS 82-07294; AVRO DAABO7-84-K-FO77; NIH 1-RO1-HL-29985-01 DARPA/ONR N0014-85-K-0807.

which of the cups should be taken away.)

There are several reasons why one might wish to give robots this same capability of reasoning in terms of generic objects. First, it is much less time- and space-consuming to model the concept of a screwdriver, than it is to model every screwdriver which the robot will encounter during the execution of its task. Second, it makes the robot system more robust. Slight deviations from the modeled object should not cause error conditions, yet deviations – such as might be caused by a misshapen tool or a malfunctioning sensor – often throw off the entire matching mechanism of a geometric model-based system. A less rigidly structured model is more robust to deviations, since it is based upon qualitative rather than quantitive measures. Finally, such a capability endows the robot system with greater flexibilty – the introduction of a new type of screwdriver into the task would not require new programming, as the robot would already be familiar with the concept of "screwdrivers."

Of course, there are many questions associated with the task of providing robots with the ability to deal with generic objects. How are such classes of objects defined, for instance? How are they reasoned about? What it the best mechanism for modelling generic objects? And how does perceptual/sensory data interact with this conceptual model? In this paper we address some of these questions with respect to a robotic perceptual system utilizing passive vision and active touch to recognize generic objects from the kitchen domain.

2. Category Theory

People tend to divide the world into categories. Tables and chairs are furniture, for example, while cats and dogs are animals. Using category theory, psychologists attempt to explain the formation, structure, and representation of these categories. And it is category theory – specifically the idea of basic-level categories – from which springs the concept of generic objects.

A category is a group of objects which may be considered similar. One way in which categories are related is by means of class inclusion. That is, sets of categories form a hierarchy of varying levels of abstraction. Sets at higher levels are more abstract than those at lower levels. In addition, categories at lower levels are completely included in categories at all higher levels. From this taxonomy comes the concept of *basic-level categories* [8], wherein certain levels of category hierarchies take on special psychological salience. For example, in the hierarchy animal-mammal-dog-poodle, dog would take on the role of basic-level object or category. The idea is that basic categories are the least abstract level of the hierarchy for which the overlap with other categories is minimized. For example, one can picture something that is just a dog, while it would be difficult to picture something that is just a mammal; on the other hand, objects further down in the hierarchy tend to share many attributes -- poodle and collie, for instance. In psychological terms, basic categories seem to provide the greatest clue validity, and they have been hypothesized as the most likely output of the perceptual system [3]. A

generic object may be thought of as a representation of this basic level for a given category hierarchy.

How are generic objects defined and reasoned about? One theory is that of prototypes. The basic level category is defined in terms of a set of features associated with a prototypical instance of that category. For example, a prototypical cup would have a handle, a cavity, and the capability of being drunk from. To determine if an instance is a member of the category, it is compared to the prototype for that category. It is not necessary for any of the objects in the category to have all of the defining attributes of the prototype. A similarity metric of some sort is applied to determine whether or not the object belongs to the category. It has also been suggested that parts and features, along with part configuration, are used to distinguish between basic level objects [10]. Parts and part configuration are important perceptually because they determine the underlying shape of an object. They also underlie behavior, since we tend to interact with objects at the parts level.

3. Representing Generic Objects

Since we want our robot to be able to explore, to identify, and eventually to manipulate generic objects, we must represent such objects within our system. Most previous work in object modelling for robotics has concentrated on geometric techniques. These modelling techiques use constructs such as generalized cylinders [2], bicubic splines [1], and planar polygons [4] to represent objects. Unfortunately, none of these techniques are flexible enough to allow for the wide range of variations to be found within an object category. Consider, for example, the range of shapes, sizes, rim diameters, and handles which different cups may contain. Yet we seldom have trouble identifying cups as such, and our robot shouldn't either. In addition, we would like to include other than geometric information in our object model. If, for example, we want to reason about objects for manipulation and task execution, it would be nice to be able to include in our representation such knowledge as "the handle of the cup can be grasped and used to lift it." For these purposes, the symbolic representations of Artificial Intelligence would seem to be more appropriate.

Thus our representation requires several properties. We must be able to handle the variations of generic objects. We must have spatial/geometric information for exploration. And we must have knowledge in the form of symbolic information for reasoning. Taking these requirements into account, along with the premise of category theory that people represent and reason about objects based upon features, we have chosen a feature-based model for our system. This representation consists of a hierarchy of frames and a spatial/geometric model which we call the *spatial polyhedron*.

The spatial polyhedron is conceptually similar to Koenderink's aspects [7]. The idea is that all of the infinte 2D views of a 3D object can be grouped into a finite set of equivalence classes. An aspect represents one such equivalence class for a given object. Aspects have

been used in computer vision by Ikeuchi [5]. In this work, 3D solid models were used to generate all possible aspects for an object in the form of an interpretation tree. This tree was then used for recognition in bin picking tasks.

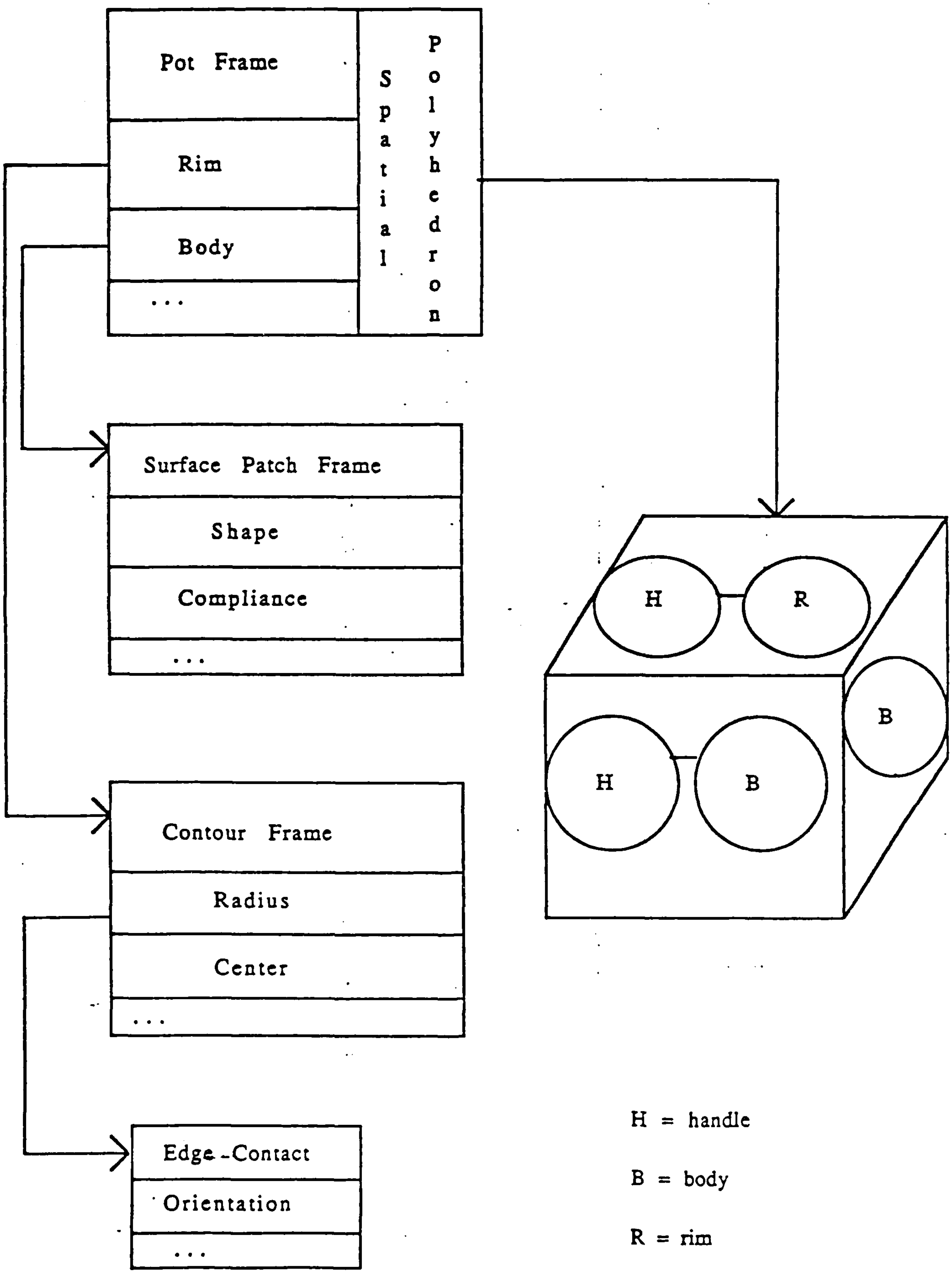
Our own approach is quite different. As we stated above, we do not want to use a geometric modelling technique to represent our generic objects. Yet we need a model which will allow us to represent the relations among the features which define such an object. In addition, we want to use this model to guide further exploration of the object -- which may contain any of a wide range of values for each defining component. For these purposes, we have devised the the spatial polyhedron. This representation may be described informally as follows. Imagine an object at the center of an n-sided polyhedron. If the object were to be viewed, or sensed, along a line normal to each face of this polyhedron, then certain components and features of the object would be viewable, while all others would not. Slight changes in attitude as the viewer moves around the object will not result in any new features coming into view. When the viewer has moved sufficiently, however, then he will be sensing the object from a different "perspective" (or face of the spatial polyhedron) and different components and features will be viewable. Thus we model an object by mapping to each face of the spatial polyhedron all of the features which we expect to be "viewable" along that face. This mapping consists of a list of these features and their appearance from the specified view. The comparison between Koenderink's aspects and the faces of the spatial polyhedron is immediate.

The remainder of our object representation consists of a hierarchy of frames. At the highest level is information about the object as a whole. Intermediate levels contain the components which define the object. The features which parameterize these components are incorporated into the spatial polyhedron. This frame representation will also carry such non-perceptual knowledge as function, ownership, etc.

We have implemented this representational paradigm for generic objects from the kitchen domain. Currently our spatial polyhedron consists of six sides for each object. For simpler objects, fewer sides might be used, while for more complex objects with larger numbers of components and features, more faces would be needed. Figure 3-1 shows a simplified version of the representation of a pot, including the spatial polyhedron. The frame hierarchy contains perceptual information about the object, while the spatial polyhedron provides spatial and relational information. So, for example, with the representation in the configuration shown, if the pot were to be sensed from above, then the rim and the handle would be encountered.

Figure 3-2 shows the prolog implementation of this representation of a pot. The integers are upper and lower bounds on enclosing volumes, radii, etc. The face clauses implement the spatial polyhedron for the object. Note that the parameters for each feature in a view are included in the representation -- we know that the handle of the pot will appear extended if

Figure 3-1: Representation of a pot.



sensed from side2, for instance.

4. Exploring Objects

We have implemented a robotic perceptual system which utilizes passive vision and active touch. The system consists of a tactile sensor mounted on a PUMA 560 robot arm and a pair of CCD cameras. Both the sensor/arm and the cameras are interfaced to a VAX 750. In Stansfield [9], we present the structure and control within this system. For the purposes of this paper, we need only give an overview of the system and its outputs.

The perceptual system is structured as a distributed hierarchy of domain specific and informationally encapsulated modules. These modules extract and identify a set of primitives and features from the object being explored. This structure is based upon Fodor's [3] theories concerning the structure of the human perception system and those of Lederman and Klatzky [6] concerning human touch. Briefly, the object to be identified is first processed visually to obtain 3-D edges and 2-D regions. Figures 4-1, 4-2, and 4-3 show the greyscale image of a pot, along with the edge and region analysis. These edges and regions are then used to invoke a set of haptic (or touch) modules which do a further exploration of the object to obtain a final set of features and components for the explored object. Figure 4-4 shows the results of this tactile exploration of the visible portions of the pot in figure 4-1.

At this point, the exploration is not model driven. The EPs are invoked based upon an initial, local, tactile exploration of the extracted visual features. But this visual data is sparse and highly inaccurate and it does not provide enough information to establish an initial arm/finger configuration. Our solution to this problem is to establish a series of predetermined "sensing planes" which are used for the intial approach toward the object. We then explore each of the visual features which has a component in the current plane. We presently approach the object from above, left, right, and front. The results, in addition to the 3D points used to generate figure 4-4, are a set of extracted features for each component of the object in each plane and a set of volumes for each visible component of the object. Figure 4-5 shows the results of exploring the pot in figure 4-1 for each plane. Note that the system does not attempt to explore a component if another component is in the way. The region labels correspond to the grey levels shown in figure 4-3.

5. Reasoning About Objects for Identification

It is immediately apparent that the results of the visually-guided exploration provide us with a structure very similar to that of our object representation – the approach planes map into the faces of the spatial polyhedron, while the volumes and object segmentation provide information to fill the slots of the frame hierarchy. Figure 5-1 shows the results of figure 4-5 in just such a form as implemented in prolog.

Figure 3-2: Prolog implementation of pot representation

```
object(one_handled_pot,50,300,80,400,200,100,
                        3,[body,part],[body,handle]).

component(one_handled_pot,body,140,140,80,
                        250,250,100,body).

component(one_handled_pot,part,50,10,10,
                200,20,20,handle).

face(one_handled_pot,2,
          [[body,contour,[rim,curved,0,[60,150,60,150]],rim],
          [handle,fpart,[large,one_extended],handle]], side1).

face(one_handled_pot,2,
         [[body,surface,[nonelastic,noncompliant,smooth,
           planar,[border,curved,0,[60,150,60,150]]],
                 bottom_surface],
         [handle,fpart,[large,one_extended],
                handle]],side2).

face(one_handled_pot,2,[
           [body,surface,[nonelastic,noncompliant,smooth,
               curved,[]],side_surface],
           [handle,fpart,[small,elongated],
                 handle]], side3).

face(one_handled_pot,1,[[body,surface,
                        [nonelastic,noncompliant,smooth,
                         curved,[]],side_surface] ], side4).

face(one_handled_pot,2,[
           [body,surface,[nonelastic,noncompliant,smooth,
                curved,[]],side_surface],
           [handle,fpart,[large,one_extended],
                     handle] ], side5).

face(one_handled_pot,2,[ [body,surface,
         [nonelastic,noncompliant,smooth,
           curved,[]],side_surface],
         [handle,fpart,[large,one_extended],
                handle] ], side6).
```

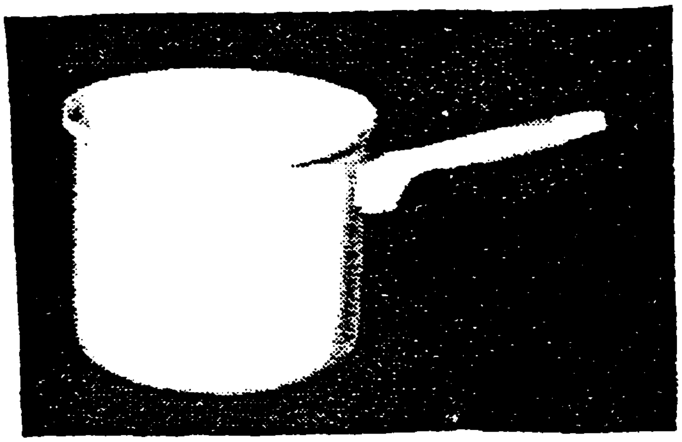


Figure 4-1: Greyscale image of a pot.

Figure 4-2: Stereo matches for the pot

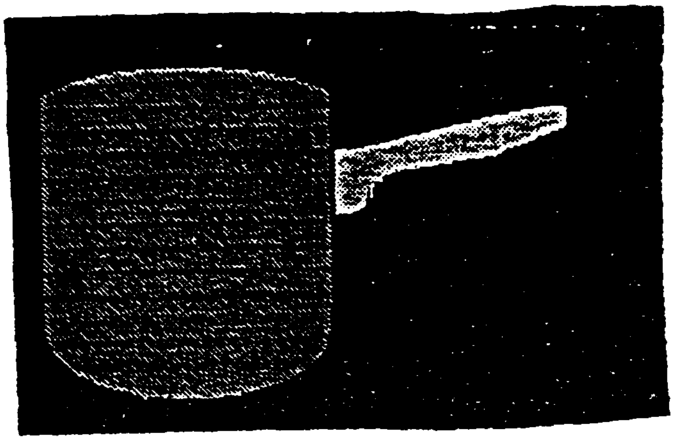


Figure 4-3: Region analysis for the pot

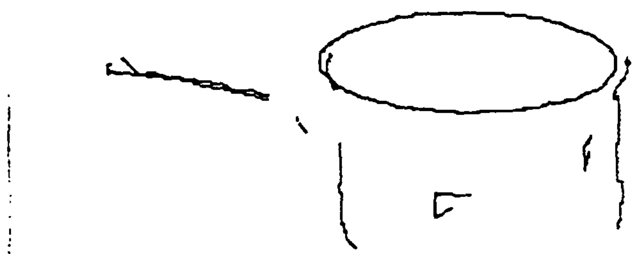


Figure 4-4: 3D results of the exploration of the pot.

Figure 4-5: Results of exploration of pot in figure 4-1.

```
view is top

region is 63                    region is -189
component is a body             component is a part
feature is contour              feature is part
contour is rim type             part is large
contour is curved               part is extended in x
radius is 96.68                 part is stubby in y

view is left

region is 63                    region is -189
component is a body             component is a part
component was not               part is small
explored haptically             part is elongated in y
reason is relational:           part is patch-like in z
-189 is left of 63

view is right

region is 63
component is a body
feature is surface patch
surface is smooth
surface is not compliant
surface is not elastic
shape is curved

view is front

region is 63                    region is -189
component is a body             component is a part
feature is surface patch        part is large
surface is smooth               part is extended in x
surface is not compliant        part is stubby in z
surface is not elastic

volumes are:

region is -189                   region is 63
xmin -478.59 xmax -387.56        xmin -703.94 xmax -512.69
ymin  102.25 ymax  102.30        ymin   36.78 ymax  225.81
zmin -144.32 zmax -144.25        zmin -278.38 zmax -156.00
```

Figure 5-1: Prolog implementation of explored pot.

```
object(obj,281,189,122,281,189,122,3,[body,part],[]).

component(obj,body,191,189,122,191,189,122,body).

component(obj,part,10,10,90,10,10,90,part).

face(obj,2,[[body,contour,[rim,curved,0,[97,97,97,97]],
                          rim_contour],
       [part,fpart,[large,one_extended],fpart]],top).

face(obj,2,[[body,surface,[unexplored],
                  surface],
             [part,fpart,[small,elongated],fpart]],left).

face(obj,1,[[body,surface,[nonelastic,noncompliant,
                   smooth,curved,[]],
                   curved_surface]],right).

face(obj,2,[[body,surface,[nonelastic,noncompliant,
                  smooth,curved,[]],
                  curved_surface],
             [part,fpart,[large,one_extended],fpart]],front).
```

The most important difference to note between the modelled pot in figure 3-2 and the data for the explored pot in figure 5-1 is that while in the model we may use cognitive lables such as handle and side surface, in the sensed data we may use only perceptual lables such as part and curved surface. This is because we have not yet matched the sensed data to an instantiated model.

The sensed object is matched against the database using a form of prototype matching. Reasoning is feature-based. The object is matched against the modelled prototypes using the extracted components, features, and their spatial relations. We require that each feature of the unknown object be present in the instantiated model, that it fit within the bounds of the upper and lower limits stored in the model, and that the relations between the instantiated and extracted features be the same. Simultaneously, the orientation of the spatial polyhedron is fixed for each matched model.

Figure 5-2 shows the results of matching the data in figure 5-1 against a database containing 19 objects. All reasoning modules are implemented in prolog. In this case, there is only one hypothesis about the object's identity, and so the system merely suggests how this hypothesis may be verified by exploring the unseen portions of the object. Information about where the features of the object are and how they should appear from these unsensed views comes directly from the instantiated spatial polyhedron.

Figure 5-2: Results of matching data in figure 5-1.

```
Object hypothesis is: one_handled_pot
matched faces are:
                    top   bottom left right front back
                    side1 side2 side3 side4 side5 side6

There is only one hypothesis, so further
exploration is unnecessary.

To verify the hypothesis, explore the back of the
object for the following:

Component is body
The explorable feature is  side_surface
It has the following characteristics:
     (surface) nonelastic noncompliant smooth curved []

Component is handle
The explorable feature is  handle
It has the following characteristics:
     (fpart) large one_extended

handle is on the left

Also explore the object from beneath for the following:

Component is body
The explorable feature is  bottom_surface
It has the following characteristics:
     (surface) nonelastic noncompliant smooth planar
        [border,curved,0,[60,150,60,150]]

Component is handle
The explorable feature is  handle
It has the following characteristics:
     (fpart) large one_extended

handle is on the left
```

6. Reasoning for Further Exploration

In the case where there are multiple hypotheses concerning the object's identity, the system generates strategies for distinguishing among them. The system reasons from the more complex hypothesis to the less complex. So, for example, it looks first for missing components, then for non-visible features of present components. The results shown in figures 6-1 - 6-4 show this method for the case of the pot in figure 4-1 turned so that the handle is occluded from the visual system.

In this case, the system does not have enough information to distinguish between the bowl and the pot hypotheses, so it determines that the handle should be looked for. The spatial

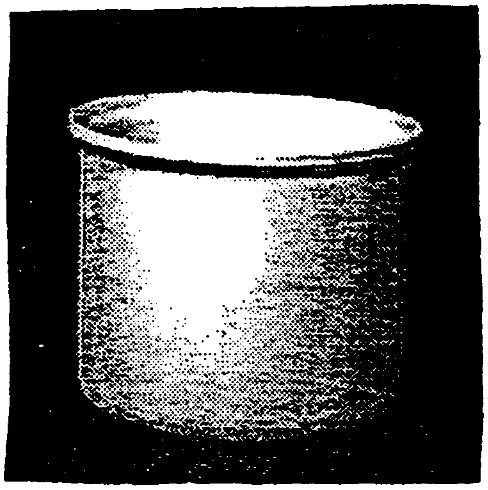


Figure 6-1: Greyscale image of a pot with handle occluded.

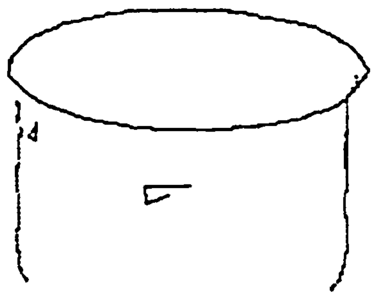


Figure 6-2: 3D results of exploring this pot.

Figure 6-3: Results of exploration and matching for this pot

```
object(obj,193,182,123,193,182,123,3,[body],[]).

component(obj,body,193,182,123,193,182,123,body).

face(obj,1,[[body,contour,[rim,curved,0,[101,101,101,101]]
            ,rim_contour]],top).

face(obj,1,[[body,surface,[nonelastic,noncompliant,smooth,
              curved,[]],curved_surface]],left).

face(obj,1,[[body,surface,[nonelastic,noncompliant,smooth,
              curved,[]],curved_surface]],right).

face(obj,1,[[body,surface,[nonelastic,noncompliant,smooth,
               curved,[]],curved_surface]],front).


Object hypothesis is: bowl
matched faces are:
                    top   bottom left  right front back
                    side1 side2  side5 side6 side3 side4

Object hypothesis is: one_handled_pot
matched faces are:
                    top   bottom left  right front back
                    side1 side2  side5 side6 side4 side3

If object is bowl then these components are missing:
none

If object is one_handled_pot then these components are missing:
handle
```

polyhedron provides information concerning the appearance of the missing component in each view for which it would be visible.

7. Handling Generic Objects

Thus far, we have shown that our system can identify objects and reason about them for further exploration and hypothesis disambiguation. In this final section, we would like to present a set of results which shows that the system is capable of handing generic objects. We have run experiments with several objects, including different plates, containers, pitchers, and bowls. If the system is to handle generic objects, then a single representation, such as that for a bowl shown in figure 7-1, must be sufficient to allow the system to identify very different types of bowls. Figures 7-2 - 7-4 show the results of the exploration and matching for a small salad bowl, while figures 7-5 - 7-7 show these results for a large mixing bowl.

As you can see, the system has generated correct hypotheses concerning the identity of

Figure 6-4: System generated strategies for further exploration.

```
To explore the object further, do the following
(Suggestions are in order of priority):

If the object is a one_handled_pot then look for
the following component(s):

Component is handle
handle is explorable from the top
From this view, the approachable feature is fpart
and it has the following characteristics: large one_extended

handle is explorable from the left
From this view, the approachable feature is fpart
and it has the following characteristics: large one_extended

handle is explorable from the right
From this view, the approachable feature is fpart
and it has the following characteristics: large one_extended

handle is explorable from the back
From this view, the approachable feature is fpart
and it has the following characteristics: small elongated

handle is explorable from the bottom
From this view, the approachable feature is fpart
and it has the following characteristics: large one_extended

If the object is a bowl then there are no missing components
Explore the object from behind to verify the following:

Component is body
The explorable feature is  side_surface
It has the following characteristics:
     (surface) nonelastic noncompliant smooth curved []

Also explore the object from beneath to verify the following:

Component is body
The explorable feature is  bottom_surface
It has the following characteristics:
     (surface) nonelastic noncompliant smooth planar
        [border,curved,0,[20,40,20,40]]
```

Figure 7-1: Representation of a bowl.

```
object(bowl,100,100,50,300,300,150,3,[body],[body]).

component(bowl,body,100,100,50,300,300,150,body).

face(bowl,1,[  [body,contour,[rim,curved,0,
                [70,150,70,150]],rim] ], side1).

face(bowl,1,  [  [body,surface,[nonelastic,noncompliant,smooth,
                   planar,[border,curved,0,[20,40,20,40]]],
                         bottom_surface]], side2).

face(bowl,1,[[body,surface,[nonelastic,noncompliant,smooth,
               curved,[]], side_surface] ], side3).

face(bowl,1,[[body,surface,[nonelastic,noncompliant,smooth,
               curved,[]], side_surface] ], side4).

face(bowl,1,[[body,surface,[nonelastic,noncompliant,smooth,
               curved,[]], side_surface] ], side5).

face(bowl,1,[[body,surface,[nonelastic,noncompliant,smooth,
               curved,[]], side_surface] ], side6).
```

both of these very different types of bowls. In the case of the mixing bowl, because of its size, the system could not distinguish between a bowl and a pot with its handle occluded, and so it has generated the second hypothesis as well. Note also that, for the salad bowl, the system has generated a correct hypothesis based upon data from the top of the object only, since it was not physically able to explore the sides.

8. Conclusion

In this paper we have introduced the concept of generic objects and presented a paradigm for representing and reasoning about them. These ideas have been implemented within the framework of a robotic perceptual system utilizing vision and touch. We discussed this system briefly and then presented the results of running experiments on several different objects. The results of these experiments show that the system is capable of identifying generic objects and of reasoning about them to generate further exploration strategies for the purpose of hypothesis disambiguation.

Figure 7-2: Greyscale image of a salad bowl.

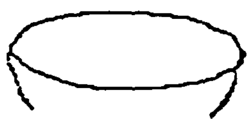


Figure 7-3: 3D results of exploring the salad bowl.

Figure 7-4: Results of exploration and matching for the salad bowl.

```
object(obj,144,140,50,144,140,50,3,[body],[]).
component(obj,body,144,140,50,144,140,50,body).
face(obj,1,[[body,contour,[rim,curved,0,
          [74,74,74,74]],rim_contour]],top).
face(obj,1,[[body,surface,[unexplored],surface]],left).
face(obj,1,[[body,surface,[unexplored],surface]],right).
face(obj,1,[[body,surface,[unexplored],surface]],front).

Object hypothesis is: bowl
matched faces are:
                    top   bottom left  right front back
                    side1 side2  side5 side6 side3 side4

There is only one hypothesis, so further exploration is unnecessary.
```

Figure 7-5: Greyscale image of a mixing bowl.

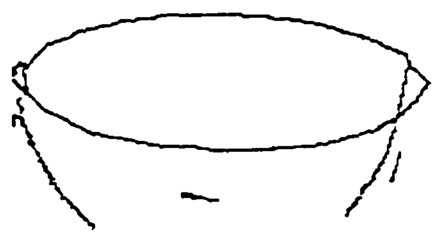


Figure 7-6: 3D results of exploring the mixing bowl.

Figure 7-7: Results of exploration and matching for the mixing bowl

```
object(obj,223,215,111,223,215,111,3,[body],[]).
component(obj,body,223,215,111,223,215,111,body).
face(obj,1,[[body,contour,[rim,curved,0,
           [107,107,107,107]],rim_contour]],top).
face(obj,1,[[body,surface,
            [nonelastic,noncompliant,smooth,curved,[]],
            curved_surface]],left).
face(obj,1,[[body,surface,
            [nonelastic,noncompliant,smooth,curved,[]],
            curved_surface]],right).
face(obj,1,[[body,surface,
            [nonelastic,noncompliant,smooth,curved,[]],
            curved_surface]],front).

Object hypothesis is: bowl
matched faces are:
                    top   bottom left  right front back
                    side1 side2  side5 side6 side3 side4

Object hypothesis is: one_handled_pot
matched faces are:
                    top   bottom left  right front back
                    side1 side2  side5 side6 side4 side3

If object is bowl then these components are missing:
none

If object is one_handled_pot then these components are missing:
handle
```

References

1. Allen, P. *Object Recognition Using Vision and Touch.* Ph.D. Th., University of Pennsylvania, December 1985.

2. Brooks, R. "Symbolic Reasoning among 3-D models and 2-D Images". *Artificial Intelligence 17* (1981), 285-349.

3. Fodor, J.. *The Modularity of Mind.* MIT Press, 1983.

4. Gaston, P. and T. Lozano-Perez. "Tactile Recognition and Localization Using Object Models: the Case of Polyhedra on a Plane". *IEEE PAMI 6*, 3 (1984), 257-266.

5. Ikeuchi, K. Precompiling a Geometric Model into an Interpretation Tree for Object Recognition in Bin-picking Tasks. Proceedings of the Image Understanding Workshop, DARPA/ISTO, 1987, pp. 321-339.

6. Lederman, S., & Klatzky, R. (1987). Hand movements: A window into haptic object recognition. Cognitive Psychology, 19, 342-368.

7. Koenderink, J. and A. van Doorn. "The Internal Representation of Solid Shape with Respect to Vision". *Biological Cybernetics 32* (1979), 211-216.

8. Rosch, E. and B. Lloyd, Eds.. *Cognition and Categorization.* Lawrence Erlbaum Assoc., 1978.

9. Stansfield, S. Visually-Aided Tactile Exploration. Proceedings of the IEEE Conference on Robotics and Automation, April, 1987.

10. Tversky, B. and K. Hemenway. "Objects, Parts, and Categories". *Journal of Experimental Psychology 113*, 2 (June 1984), 169-193.

A MEDIUM-COMPLEXITY COMPLIANT END EFFECTOR

Nathan Ulrich, Richard P. Paul
and Ruzena Bajcsy

MS-CIS-87-82
GRASP LAB 118

Department of Computer and Information Science
School of Engineering and Applied Science
University of Pennsylvania
Philadelphia, PA 19104-6389

September 1987

Acknowledgements: This work was supported in part by National Science Foundation Grants MEA-8119884, DCR-8410771, CER/DCR-MCS82-19196, INT-8514199, DMC-8517315 and DARPA/ONR grant N0014-85-K-0807.

A Medium-Complexity Compliant End Effector*

Nathan Ulrich Richard P. Paul Ruzena Bajcsy

Department of Mechanical Engineering
Department of Computer and Information Science
University of Pennsylvania
Philadelphia, PA 19104

Abstract

Recent interest in end effector design has not yet resulted in a versatile yet simple mechanism appropriate for a wide range of manipulation tasks. The design of a novel end effector under development at the University of Pennsylvania is explained in detail in this paper. The rationale supporting this mechanism is explored, its geometry is described, experimental results from the first prototype are shown, and some ideas for future work are presented.

Introduction

In recent years there has been a great deal of attention focused on the design of end effectors. Progress in grasping research, active sensing, assembly, and

*Supported by NSF grants MEA-8119884, DCR-8410771, CER/DCR-8219196, INT-8514199, DMC-8517315, and DARPA/ONR grant N0014-85-K-0807. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the supporting agencies.

prototype construction has created a need for a versatile, robust, and economical mechanical hand that can be used for experimentation. Although many designs have been proposed and several prototypes built, a comprehensive effort which combines the desire for performance with the reality of application has yet to be undertaken. As a result, no single device is in common use.

Most previous end effector designs fall into two categories: complex "hands" or simple grippers. Notable in the first class are the Utah/MIT Dextrous Hand [1] and the Salisbury hand [2]. They incorporate a large number of degrees of freedom (degrees of freedom) into a complex multi-fingered hand design which imitates the human hand in speed, dexterity, and versatility. The resulting performance is impressive, but the increased complexity precludes simple planning procedures. The simple grippers do not have this problem—they are generally one or two degrees of freedom and are powered by means of remote pneumatic or self-contained electric actuators. They pay for this simplicity by being limited in application, usually specialized for one type of task.

We feel that what is needed is a *medium-complexity* end effector: a device that combines the simplicity characteristic of the simple grippers with some of the versatility of the complex hands.

Design Philosophy

The design of any tool requires a precise definition of its intended use. It is important to not only decide what tasks a robotic end effector needs to be able to perform, but to also determine the limits of its performance. Previous hand designs have used the human hand as a so-called "existence proof" of the appropriateness of such a geometry. Since our hands are capable of many varied

tasks, any mechanical end effector which duplicated the human hand would also be capable of these tasks. But this is not sufficient reason for an anthropomorphic geometry. The design of an end effector should be pursued in the same way as any other design; establish the criteria for its performance and synthesize a mechanism which satisfies these goals. For our specific research environment, the end effector is required to machine and assemble parts, handle many different sizes and shapes of objects, and perform exploratory and sensing tasks—it does not need to be able to perform tasks outside of this environment. While the human hand seems to be ideal for performing the wide range of tasks required of a person—from playing basketball to changing diapers to driving nails—it is not necessarily the perfect tool for the specific areas in which robotic research is now concentrated. Witness the number of tools to assist the human hand found in a machine shop. It should be possible to design an end effector that is more suited than the human hand for such an environment.

Design Criteria

The Medium-complexity Compliant End Effector (McCEE) is designed primarily for three research areas: active sensing, assembly (and disassembly), and grasping.[1] Although these subjects encompass a wide range of criteria, we feel that they overlap sufficiently for the use of one basic end effector design.

Grasping research requires a versatile mechanism that allows application of theoretical methods to experimental situations. The state of the art at this point demands a more flexible tool than the simple grippers commonly used,

[1] Research in the application of this design to prosthetics is continuing, but is beyond the scope of this paper.

but it is extremely important that the complexity of the end effector be limited. Since theoretical principles cannot support a complex (e.g. 9 or more degrees of freedom) model of grasping in three dimensions, we feel that a medium-complexity device is most appropriate at this time. The simplicity of planning, movement, and control associated with fewer degrees of freedom is an important consideration—such a tool would be more accessible to the researcher. However, it is important to note that 9 degrees of freedom is the minimum necessary to allow arbitrary positioning of three fingertips in space. For this reason, our design will concentrate on *enveloping* grasps; those that rely on the palmar surfaces of the inside of the fingers and the palm to constrain an object, as opposed to fingertip manipulation utilizing friction and fingertip contacts[3]. An extension of the two degree of freedom grippers is necessary, but in interest of utility, we would like to limit our end effector design to three or four degrees of freedom.

Although recent advances in vision and other passive sensing techniques have resulted in increased reliability and information gathering ability, it has been shown that the use of active sensing is necessary to adequately define the shape and orientation of an object[4][5][6]. In addition, psychological research has defined a number of "exploratory procedures" that can be used to collect such characteristics of an object such as texture, hardness, thermal conductivity, and shape[7]. Such sensing will allow us to classify an object or verify a hypothesis; an exact description is essential to allow us to perform manipulation in an assembly operation or to support grasping experimentation. Therefore, the end effector will need to serve as a platform for a number of specialized sensors necessary for this work. It is necessary that a sensor package be incorporated in

the design of the end effector, but that the end effector be sufficiently versatile to accomodate changes in sensor type and application. The primary sensors—those integral to the design—provide position, tactile, force, and moment information on contact surfaces. But the design must also consider easy mounting and dismounting of other more exotic sensors (thermal and electrical conductivity, proximity, specialized textural, etc.).

Assembly of parts and objects is an important area of robotics research because of its relevance to industrial applications. However, assembly tasks performed by robots today are limited to rigid, structured operations which usually require complex jigs and parts-feeding devices. Any appreciable uncertainty in such an operation cannot be accomodated. This is essentially automation and *not* robotics. At a certain level of production capacity, such automation becomes cost effective. However, below this critical level, human workers are necessary to supplement any generic automatic devices in use. A true robotic assembly operation would combine grasping and sensing with computational sophistication, and would be able to tolerate much larger errors in positioning and description. Necessary to such an operation, however, are one or more versatile end effectors that are suited for both a wide range of grasps and a variety of sensors. Such a device should be able to handle both parts and tools, as well as possessing the sensor sophistication to recognize and differentiate objects. But even with these capabilities, an assembly operation still requires a model and procedure to follow. Previous research has used human-based techniques to synthesize assembly algorithms. However, the strengths and weaknesses of a robotic system are inherently very different from those found in humans. By taking an object apart, finding seams, joints, and fasteners, such a system could determine the

best way for a robot to reassemble the object. The ability to perform effectively in such a disassembly operation is an important criterion for our end effector design.

A number of criteria for the design of an end effector that could perform the operations suggested above are related to convenience and utility. The mechanism would ideally be self-contained; discrete from the manipulator and able to be mounted and dismounted quickly and easily to facilitate adjustment and repair. A compact, sleek design integrating all cabling, sensing, and actuation is important, but since it will be a research tool, the mechanical design should be accessible, allowing changes in structure and operation without radical reconstruction or redesign. The use of the end effector to learn about objects necessitates it use as a platform for many types of sensors. All of these sensors do not initially need to be built-in, but the design must be able to accomodate their use. The end effector should, ideally, satisfy the research imperatives described previously while attaining these objectives as well.

Supporting Research

Many researchers have attempted to classify the grasps required by a robotic end effector. Schlesinger defined six prehension types used by humans in his work[8], and Cutkosky and Wright further defined the grasps used by a machinist at work[9]. Although other, different, classifications have been used (see [10] for a complete grasp taxonomy), we find these two sets of descriptive labels most appropriate for our applications. The grasps required by assembly, disassembly, prototype construction, and grasping research are contained within these types, represented graphically in Figures 1 and 2.

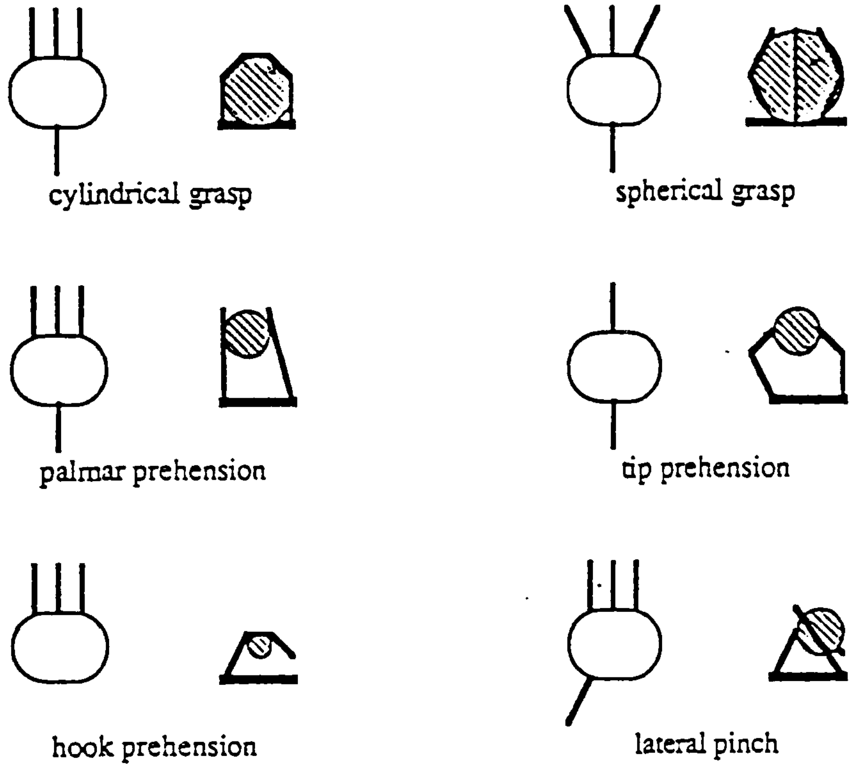


Figure 1: Schlesinger's prehension types

While the actual apprehension of an object with a robotic end effector can be modeled using the above classifications, the use of the device as a tool for active sensing requires expansion of these models. Although a great deal of haptic (kinesthetic plus tactile) information can be gained by simply holding an object, the exploratory procedures described by Klatzky and Lederman require other sensory methods. Figure 3, adapted from [7], shows the properties that we need to obtain by active sensing and the necessary actions of the end effector to determine these properties. In order to perform these movements with an end effector, we need several abilities. First, we need to be able to use the end effector with one finger extended as a probe. This will allow us to perform the exploratory procedures to test for texture, hardness, temperature, and will allow us to determine the shape of the object by means of the procedures suggested by Allen [5] and Stansfield [6]; i.e. determine surfaces, cavities, holes and

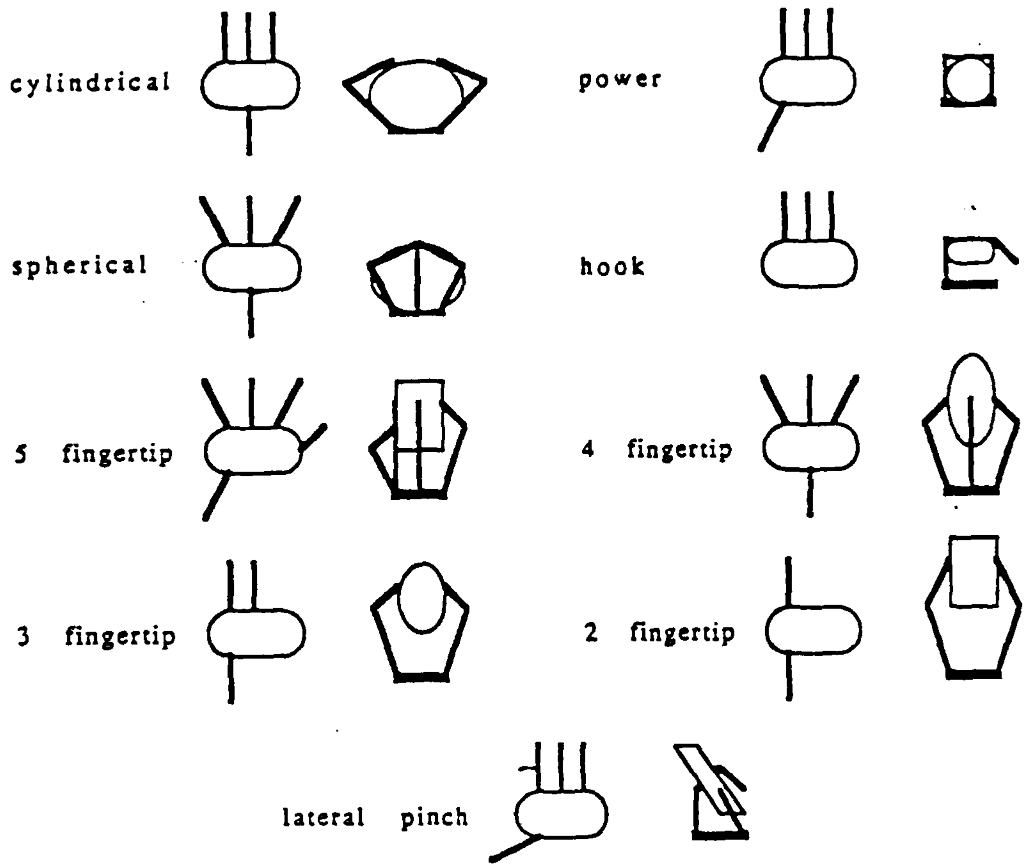


Figure 2: Cutkosky and Wright's manufacturing grips

	Properties	Hand Movements
Surface Properties	Texture Hardness Temperature Weight	Lateral Motion Pressure Static Contact Unsupported Holding
Structural Properties	(Weight) Global Shape Exact Shape Volume	(Unsupported Holding) Enclosure, Contour Following Contour Following Enclosure

Figure 3: Classification of properties and exploratory procedures

contours. In order to accomplish these tasks, this finger would need tactile sensing capability, force and position sensing, and also specialized temperature sensors.

The end effector must also be able to enclose an object within its grasp and lift it free of support. This will allow us to determine the weight, shape, and volume of the object. Such a function requires similar properties as those required by other aspects of our goals, but also requires precise sensing of the object within the grasp. A determination of an object's properties by means of the exploratory procedures described above is essential to an accurate classification of the object; such a classification is necessary for success in the assembly, disassembly, and prototype construction workplaces described previously. It follows, then, that in order for an end effector to be useful in these task-oriented environments, it must also be a efficient tool for active sensing.

Mechanical Configuration

The shape of the end effector design was determined by the need to achieve wide versatility with as few degrees of freedom as possible. We found that in order

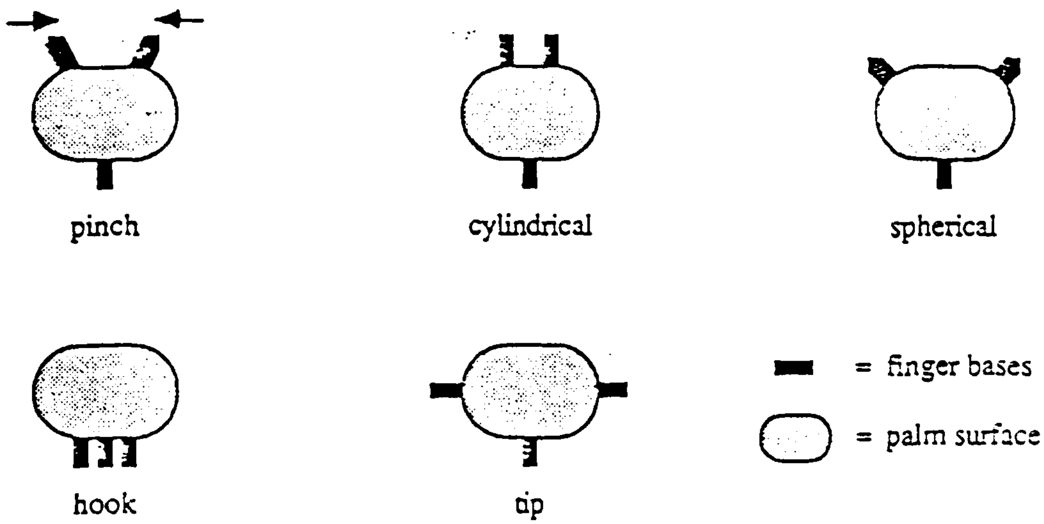


Figure 4: The five grasping modes of McCEE

to obtain the grasping and sensing configurations necessary for our research, we needed an end effector with at least four degrees of freedom. The actual mechanical geometry is separated into two parts: the shape of the palm and its relationship to the fingers, and the finger design.

The palm/finger relationship consists of a one degree of freedom movement of the fingers around the palm. Skinner proposed a similar movement of the fingers, but his design did not incorporate the palm into the grasping arrangement[11]. We wish the palm to be an important tool in the manipulation of objects. Not only can the palm be used as a base against which to hold objects, as a tool to perform pushing operations on objects, but also (with tactile sensors) as a information-gathering instrument which will allow "footprints" of objects to be obtained. By separating the centers of rotation of the fingers, we obtain a number of grasping configurations. Figure 4 shows these different modes. One finger (which, although not precise biologically, we call the thumb) has its base fixed with respect to the palm, while the other two move synchronously around two different axes. The resulting scheme allows a very wide range of grasping

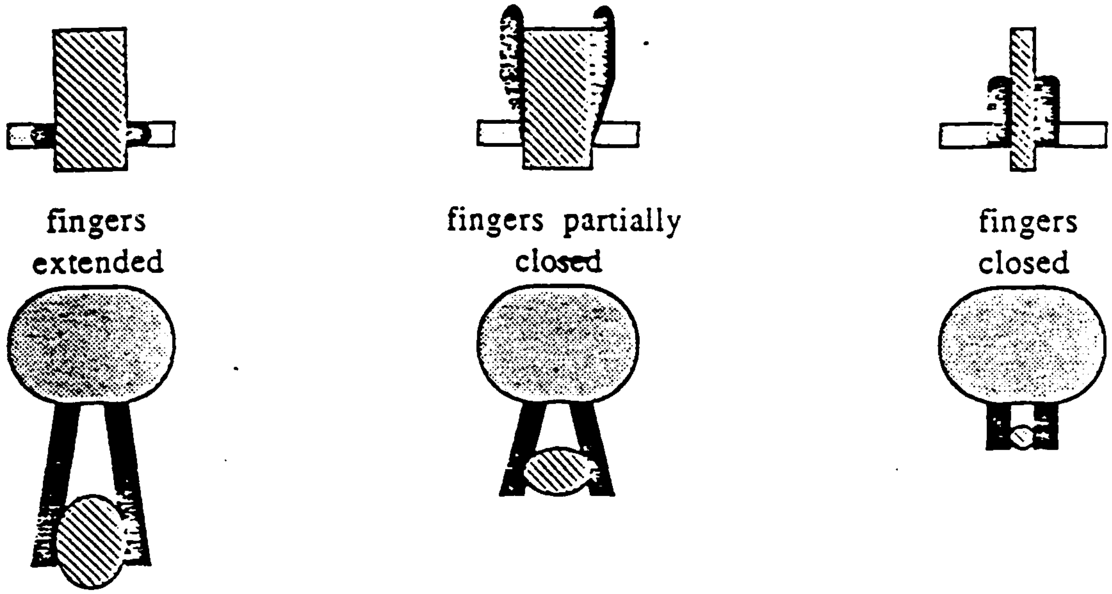


Figure 5: Variations of the pinch grasping mode

types and, in addition, yields a pinching grasp between the two fingers similar to that used by amputees who use a split hook. Another advantage to this configuration is that the palmar surfaces of the fingers are always facing directly inwards—simplifying the sensing of an object within a grasp—in contrast to the human hand, where the lateral movement of the fingers does not allow this. The five grasping modes are described below with their parallels in Schlesinger's and Cutkosky and Wright's work defined as well:

The **pinch** grip occurs when the two movable fingers are brought together on the opposite side of the palm from the thumb. The inside of these two fingers are lined with rubber, which allows for friction grasping of small objects. This is primarily a precision grasp, used for picking up small, delicate objects. It is similar to the lateral pinch grasp described by both Schlesinger and Cutkosky and Wright. In addition, some operations which are usually performed by Schlesinger's tip prehension and Cutkosky and Wright's two-finger precision grasp can be achieved in this configuration. The flexibility of this grasp is enhanced by the ability to change its nature by changing the angle of the fingers. In Figure 5, this technique is illustrated. This grasp is very similar

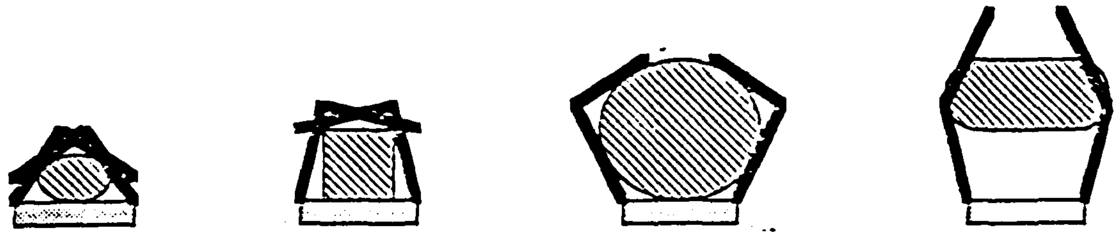


Figure 6: Variations in the cylindrical grasping mode

to the precision grasp used by amputees who have been fitted with a split hook prosthesis. In this case, a cylindrical groove between the halves of the hook allow for stable grasping of a pencil or similar small cylindrical objects. Such an implementation in the robotic end effector could prove useful.

The **cylindrical** grasp, when the two fingers are opposite the thumb, is analogous to Schlesinger's cylindrical grasp and Cutkosky and Wright's cylindrical power and precision grips. This mode allows for the apprehension of a wide range of shapes and sizes, from small cylindrical objects to larger rectangular box-shaped objects (see Figure 6). In addition, this mode allows a version of the lateral pinch grasp, when an object is held between the three fingertips. The attractiveness of this grasp lies in its strength. Since the palmar surfaces of all three fingers are holding the object against the palm, objects are held very securely.

The **spherical** grasp, with the three fingers roughly 120 degrees apart, is similar to Schlesinger's spherical grasp and Cutkosky and Wright's spherical power and 3-finger, 4-finger, and 5-finger precision grasps. In a power grasp, the palmar surfaces of the fingers are used to hold a spherical object against the palm, while in a precision grip, the three fingertips form a three-sided fingertip

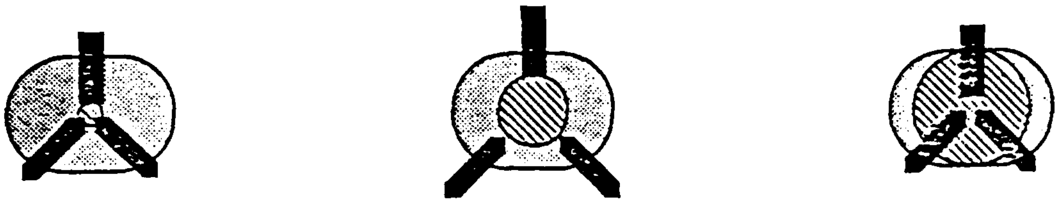


Figure 7: Variations of the spherical grasp

grasp which is similar to the chuck on a drill. In Figure 7, the application of this grasp to various objects is shown.

When the two fingers are rotated until they are opposite each other, they can be used in a **tip** grasping mode. This is exactly the tip prehension described by Schlesinger and the 2-finger precision grip described by Cutkosky and Wright. Although this grasp relies primarily on friction for stability, it can be useful in apprehending objects that are ackwardly placed or for manipulating objects securely held in some manner. The pinch grasp provides a more stable grasp of most small objects.

The **hook** mode of grasping uses all three fingers located together on one side of the palm. This allows for two types of grasping: a passive grip on a handle or similar structure where the fingers act as a hook, or an active grasp where all three fingers hold a large object against the palm. This is a grasp that could be used to lift one side of a large flat object (in cooperation with another hand) where the size of the object precludes an enveloping grasp. Figure 8 shows these uses.

Although these modes provide wide versatility in grasping, an equally flexible finger design is necessary in order to fulfill our design objectives. A finger of fixed shape pivoting around the edge of the palm would provide only limited capability. Although it could hold many objects, such a finger could only perfectly

Figure 8: Variations of the hook grasp

Figure 9: Variations in finger shape with changes in object shape

grasp a small number of objects with optimum contact points corresponding to its fixed shape. In Figure 9, we show how ideal finger shape varies with object geometry. We would like to have a finger which could change its geometry in response to the shape of the object. A multi-jointed finger such as those found on the Utah/MIT DH [1] and in the Salisbury hand [2] can comply to the object shape by integration of sensor feedback and position control. However, these fingers have 3 or 4 degrees of freedom. We need a finger which can achieve this same function without the control and actuation complexity associated with these added degrees of freedom.

The author originally proposed such a finger design in the Compliant Articulated Mechanical Manipulator (CAMM) [12], which incorporated a four-joint finger with two degrees of freedom. We have modified the design to yield a two-jointed one degree-of-freedom compliant finger design. The single degree

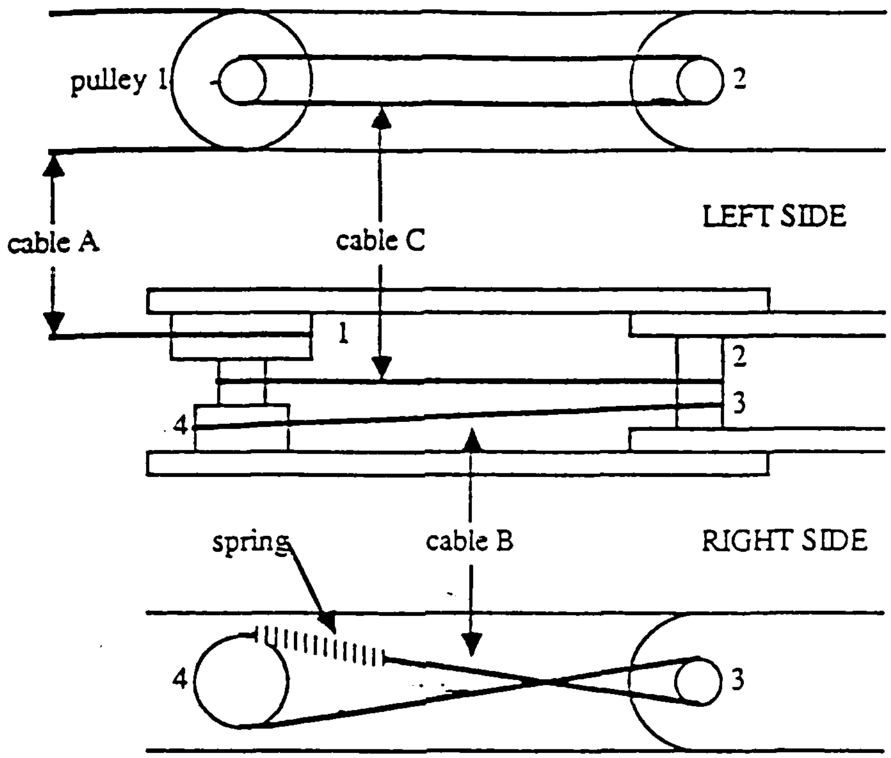


Figure 10: Schematic representation of actuation linkages

of freedom satisfies our need for simplicity, yet allows flexibility in object apprehension. Figure 10 shows a schematic of the linkages involved. This finger will *passively* shape itself to an object without the use of control computation or sensor feedback. The finger incorporates a spring in its linkage to provide compliance in one direction; this allows the second joint of the finger to continue to rotate once the first joint contacts an object. However, no matter how much the joints rotate independently, the finger will not comply in opening; that is, it will always maintain pressure on the object dependent only on the torque produced by the actuator. The compliance is implemented in the linkage contained on the right side of the finger, while at the same time the drive linkage on the left side of the finger actuates the finger and transfers gripping force. For a more detailed description of this finger and its kinematics, see [13].

Experimentation

It is common for a design to look good in theory and on paper, but to prove disappointing in implementation. To prevent the investment of time and money into a electrically-actuated, computer-controlled design that might prove useless, we decided to build a prototype of our design which would use movement of an experimenter's fingers to actuate the fingers of the end effector. This device was in essence a manual teleoperated end effector. This allowed us to test our ideas very quickly, utilizing the experimenter's brain as a control system, and his body as the actuator. It was in experimentation with this device that the actual design presented here was developed. This prototype was simple and inexpensive to build and allowed quick modification. In combination with prototypes of the finger design, we were able to finalize the design with little effort.

In the process of our experimentation, we found the device very useful; that all of the grasps necessary for enveloping grasps and tool handling were possible, and that the actions necessary for assembly and disassembly could be achieved. However, the device does have limitations. As anticipated, the design is more suited to enveloping grasps and handling large tools. Associated with the low number of degrees of freedom is a loss of dexterity in small parts manipulation. Although such objects can be grasped securely, movement of the objects within the grasp requires interaction with a table surface or another hand. We do not find this a serious fault for our work, since the use of two hands for assembly tasks is probably necessary anyway.

Conclusion

We have presented the basis of a medium-complexity compliant end effector design. The end result of our identification of a gap in end effector development has led to a four degree of freedom flexible end effector design that is especially suited for work in active sensing, assembly and disassembly, and grasping. We have attempted to support the rationale for this design on fundamental good engineering practice as well as on previous research. There are obviously many details of the design which have not been described here, but an electrically-actuated self-contained end effector for use on the end of a robotic manipulator is under construction. Use of this device will allow expansion of present research topics and allow for experimentation in new areas related to robotic manipulation.

References

[1] Jacobsen, S.C., E.K. Iversen, D.F. Knutti, R.T. Johnson, and K.B. Biggers, "Design of the Utah/MIT Dextrous Hand," *Proceedings of the IEEE Conference on Robotics and Automation*, San Francisco, April 1986.

[2] Salisbury, J.K., "Kinematic and Force Analysis of Articulated Hands," *Ph.D. Thesis, Stanford University*, July 1982.

[3] Trinkle, Jeffrey C., "The Mechanics and Planning of Enveloping Grasps," *Ph.D. dissertation*, University of Pennsylvania, June 1987.

[4] Bajcsy, R., "What can we learn from one finger experiments?" in M. Brady and R. Paul, editors, *The First International Symposium on Robotics Research*, MIT Press, Cambridge, 1984.

[5] Allen, P.K., "Object Recognition Using Vision and Touch," *Ph.D. dissertation*, University of Pennsylvania, September 1985.

[6] Stansfield, S.A., "Primitives, Features, and Exploratory Procedures: Building a Robot Tactile Perception System," *Proceedings of the IEEE International Conference on Robotics and Automation*, San Francisco, 1986.

[7] Lederman, S., & Klatzky, R. (1987). Hand movements: A window into haptic object recognition. Cognitive Psychology, 19, 342-368.

[8] Schlesinger, G., *Der Mechanische Aufbau der kunstlichen Glieder*, Part II of *Ersatzglieder und Arbeitshilfen*, Springer Verlag, Berlin, 1919.

[9] Cutkosky, Mark R. and Paul K. Wright, "Modeling Manufacturing Grips and Correlations with the Design of Robotic Hands," *Proceedings of the IEEE Conference on Robotics and Automation*, San Francisco, 1986.

[10] Iberall, Thea, "The Nature of Human Prehension: Three Dextrous Hands in One," *Proceedings of the IEEE Conference on Robotics and Automation*, 1987.

[11] Skinner, F., "Design of a multiple prehension manipulator system," *ASME publication 74-DET-25*, October 1974, and *Mechanical Engineering*, September 1975, pp. 30-37.

[12] Ulrich, Nathan, "The Compliant Articulated Mechanical Manipulator," University of Pennsylvania, GraspLab Memo, April 1985.

[13] Ulrich, Nathan, "A Two-Jointed One Degree-of-Freedom Compliant Finger," University of Pennsylvania, September 1987.

END
DATE
FILMED
7-88
DTIC